

AFRL-RI-RS-TN-2008-4
Final Technical Note
July 2008



PROACTIVE INTELLIGENCE (PAINT)

Northrop Grumman Systems Corporation

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

STINFO COPY

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE
ROME RESEARCH SITE
ROME, NEW YORK

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the Air Force Research Laboratory Public Affairs Office and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (<http://www.dtic.mil>).

AFRL-RI-RS-TN-2008-4 HAS BEEN REVIEWED AND IS APPROVED FOR
PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE DIRECTOR:

/s/

EDWARD L. DEPALMA
Work Unit Manager

/s/

JOSEPH CAMERA, Chief
Information & Intelligence Exploitation Division
Information Directorate

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE*Form Approved*
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Service, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE (DD-MM-YYYY) JUL 2008		2. REPORT TYPE Final		3. DATES COVERED (From - To) Sep 07- Dec 07	
4. TITLE AND SUBTITLE PROACTIVE INTELLIGENCE (PAINT)				5a. CONTRACT NUMBER FA8750-07-C-0171	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Manisha Singh				5d. PROJECT NUMBER PAIN	
				5e. TASK NUMBER 00	
				5f. WORK UNIT NUMBER 11	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Northrop Grumman Systems Corp. 1580A Nursery Road Linthicum Heights, MD 21090-2202				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) AFRL/RIED 525 Brooks Rd Rome NY 13441-4505				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSORING/MONITORING AGENCY REPORT NUMBER AFRL-RI-RS-TN-2008-4	
12. DISTRIBUTION AVAILABILITY STATEMENT APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED. PA# WPAFB 08-3820.					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT This report documents a Northrop Grumman contract that was awarded under the IARPA Proactive Intelligence (PAINT) program to develop a modeling technology to support the Intelligence Community in avoiding strategic national security surprises. The technology addressed the problems of identifying, predicting and managing adversarial coalitions of those intolerant of US interests. This would be achieved through the use of social science, decision science and data management techniques. The Denied area Ethnographic Environment Model (DEEM) would provide a description of current threatening situations, predictions of potential future threatening situations, rapid model implementation and interfaces for integration into a Virtual World/Dynamic Gaming Environment. Due to budget constraints, this effort was terminated at an early stage so the final result was not realized. This technical note details the preparations for the work and the steps that were taken during the short duration of the contract.					
15. SUBJECT TERMS Simulation Models, Proactive Intelligence, Sigmoid Utility, Agent-Based Model, Fractional Graph Model, Standard Ethnographic Sample					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 7	19a. NAME OF RESPONSIBLE PERSON Edward L. DePalma
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U			19b. TELEPHONE NUMBER (Include area code) 315-330-3069

Program Status

This Final Technical Note covers the period of 01 September 2007 through 21 December, 2007, and in accordance with contract FA8750-07-C-0171 agreement between Air Force Research Laboratory and Northrop Grumman Electronic Systems (NGES).

Northrop Grumman held a telecom with Mr. Daniel Daskiewicz (AFRL) on the 21st of December 2007, informing Northrop Grumman that due to the reorganization of IARPA, not program performance, AFRL was forced to terminate some efforts and scale back others. Northrop Grumman was one of the contractors whose contract was terminated. There were significant accomplishments on the 3 month Northrop Grumman PAINT effort. The following summarizes the work accomplished between the reporting periods of contract initiation to contract closeout.

The team of Northrop Grumman and Purdue University was awarded a contract on the 30th of August 2007. A program kick-off was held with Mr. Peter Brooks in Washington D.C on the 4th of September. A kick-off (workshop) was scheduled and attended on the 5th and 6th of September 2007 by all PAINT contractors. At the kick-off a challenge problem was discussed in breakout groups, as well as homework assignments were provided to each contractor. As a result of the kick-off, the Northrop Grumman team consisting of Purdue University has been focusing on the following general categories of tasks:

- Collaborating on the PAINT “homework” assignments
- Transitioning pre-existing code to PAINT specifications
- Researching the challenge problem.

The team participated in the 2-day kickoff meeting (workshop), established other team members for the homework assignment (Lockheed Martin, Least Squares and LCC) and had been continuously collaborating with the other PAINT contractors on the three agreed upon quad charts. In order to fulfill the requirements of the quads/homework assignment, the NGES team had to revise our PAINT strategy in order to fit within the larger PAINT project framework. Our 1-year schedule was revised.

The NGES team had begun to transition pre-existing code for simulating societies, producing risk sensitive behavior and evolving coalitions and alliances. This would have been an on-going process; however we were able to train a new programmer on this task and inserted extensive documentation into the code and were able to formulate a staged plan for transitioning the code to perform PAINT specifications.

Preliminary research had also begun on the challenge problem including planning how the challenge problem agents and their network connections will be represented in our simulation, as well as gathering basic data relevant to the challenge problem.

The NGES Team had also commenced its review of both format and content of the "nano" corpus and “Factiva” data from LCC. We also made assessments of the BAE data that had been posted to TWiki to scope our specific demonstration. The reviews allowed us to refine our data ingest requirements and design. Parser for Factiva data has been implemented and now tested. We have identified and now implemented critical web service components. We had implemented the affinity propagation-based clustering algorithm for use in our demonstration.

Efficient, scalable clustering implementation had been developed and tested. We had numerous discussions on the content and were considering this data in light of our existing ontology's and those anticipated from the challenge problem. The data provisioning components common to both the SOW requirements and those implied by the quad chart activities were formalized in anticipation of their independent, local integration as well as their integration into the larger, composite PAINT demonstration system.

The data provisioning components and architecture design common to both the SOW requirements and those implied by the quad chart activities were refactored against interfaces, formalized and then completed in anticipation of their independent, local integration as well as their integration into the larger, composite PAINT demonstration system. We evaluated previous code developed for research and continued to work on the transitioning the code in order to develop it for a leadership model. We extracted, analyzed, and then altered the code to maintain its basic functionality. We completed empirical research of databases, Internet sources, and academic and government literature in order to construct the simulations required for the PAINT tasks. Preliminary research conducted on data sources, the construction of an input database and exploring relevant fractional graph theoretic methods for use on our PAINT project had begun. In addition, research on the leadership of various political factions in our original SOW problem was also completed. Preliminary simulations were run to verify the execution of the codes we were developing.

Concentration was on building elements into the code that are unique to the PAINT task, including; programming a leadership model (developing appropriate data input structures, simulation environment, functionality) for representing the PAINT challenge problem, and programming a leadership model (developing appropriate data input structures, simulation environment, functionality). We wanted ensure that our work met all the original contracted work in accordance with our statement of work.

Prior to the cancellation notification, we continued to work on common tasks between the original statement of work and the quad chart assignments. Data review has been completed and preparations were in place for the initial design and implementation for the demonstration.

Program Work Summary

This is a summary of the work performed under the PAINT contract. None of the algorithms identified below were developed under the PAINT contract.

The following is the work that was completed under the PAINT contract:

- PAINT Kickoff Meetings 5-6 September
- Assessment and Preliminary Design of PAINT Challenge Problem
- New partnership affiliation meetings and telecons
- Development of new plan for new PAINT Challenge Problem
- Re-planning for in and out of scope tasks
- System Design
- System Configuration
- Extractors Analysis
- Tokenizer, Stopper, Stemmer Analysis
- POS (Part of Speech) Tagger Analysis

- Literal Term Search analysis and design
- WordNet analysis and design
- RSS/ATOM (Really Simple Syndication) Aggregate analysis and design
- A functioning simulation model of PAINT challenge problem scenario
- A functioning simulation model of a leadership model originally proposed in the Northrop Grumman/IPFW (Indiana Purdue University of Fort Wayne) proposal
- Preliminary analysis of leadership model simulations

The following work was in progress at the time that the PAINT contract was terminated. This work involved the initial analysis and suitability assessment of use of various previously developed algorithms for the PAINT project. None of the algorithms identified below were developed under the PAINT contract. An associated percent complete has been assigned for each activity:

- Web Front End(s) 25%
- ConAn (Concept Annealing) Search 50%
- Clustering 50%
- Feature Extractor(s) 25%
- Multinomial Distance Metric 50%
- Semantic Term Search 25%
- Visualization 25%
- IR Information Retrieval 50%
- Leadership Model 50%

In addition to the completion of the above work, the following work was yet to be executed under the PAINT project.

- Final design leadership model
- Integration and testing of leadership model
- Analysis of leadership model results
- Demonstration of leadership model
- Morphology (Inflection)
- SVM (Support Vector Machine)
- Term Semantic Similarity
- Syntax
- Integration
- Testing
- Demonstration
- Final Technical Note

Meeting and presentations held between 01 September 2007 to 21 December 2007:

4 September 2007- Kick-off meeting with Dr. Peter Brooks and Ed DePalma of AFRL

05-06 September 2007-Kick-off and Workshop Presentations

14 September 2007- Kicked- Off Homework Assignment with other PAINT Contractors

18 September 2007- Finalization of Quad Charts with other PAINT Contractors

1 November 2007- Initiated a telecon with Mr. Ed DePalma and Dan Daskiewicz of AFRL to discuss the DD254 as well as the additional cost associated with collaboration and homework assignments (quad charts). A spreadsheet was submitted on 01 November 2007, reflecting the additional workscope and hours associated with those tasks for the group assignment (homework/quad chart).

14 November 2007: Telecon with Lockheed Martin, MIT, and Berkeley/LCC was conducted to have discussions on the “Pathway Disambiguation Experiment” (Experiment 2) and thoughts for moving forward were formulated. The PAINT Integration plan was reviewed in detail and discussions on paths forward were investigated.

15 November 2007: Telecon with Berkeley/LCC was conducted. The PAINT Integration plan was reviewed again and discussions on the “Proactive Knowledge Aggregator” and the pathways to move forward were finalized.

15 November 2007- Telecon with PAINT customers and participants were held to discuss the “MultiMethod Expert Validation.” Participants include LSS, HNC/FI, MIT, ISI/CREATE and NGC. This telecon clarified the PAINT integration vision and plan. Each demonstration and its purposes were discussed including the Year 1 PAINT program outcomes.

Deliverables submitted to AFRL to date include the following:

PAINT Presentations on TWIKI, Quad Charts on TWIKI, Monthly Technical and Financial Status Reports.